DOCUMENT RESUME

ED 347 748 EC 301 364

AUTHOR Rosenkoetter, Sharon E.; Wanska, Susan K.
TITLE Best Practice in Preschool Screening.
PUB DATE 14 Apr 92
NOTE 32p.; Paper presented at the Annual Convention of the Council for Exceptional Children (70th, Baltimore, MD, April 13-17, 1992).
PUB TYPE Guides - Non-Classroom Use (055) -- Reference Materials - General (130)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Definitions; *Disabilities; Educational Policy; *Educational Practices; Educational Quality; *Handicap Identification; Models; *Preschool Children; Preschool Education; Program Development; *Screening Tests; Student Characteristics; Test Validity

ABSTRACT

This brief guide outlines best practices in preschool screening for the presence of possible disabilities. It covers: a definition of screening, its history, the rationale for screening, requirements for effective screening measures (reliability, validity, fairness, utility), models for preschool screening, evaluation of screening procedures, and some things that screening isn't (e.g., the basis for decisions about eligibility or placement). The bulk of the document consists of the following attachments: (1) a classification analysis to determine a procedure's validity for individual decision-making; (2) characteristics of young children which influence screening outcomes; (3) guidelines for best practice in preschool screening; (4) a summary of selected screening instruments; (5) 1991 recommendations from the U.S. Department of Education from "Preparing Young Children for Success: Guideposts for Achieving Our First National Goal--An America 2000 Education Strategy"; (6) a summary of a longitudinal study of prekindergarten transition screening involving teacher report and traditional testing; and (7) a teacher checklist. (47 references) (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED347748

BEST PRACTICE IN PRESCHOOL SCREENING

Council for Exceptional Children, Baltimore, MD
April 14, 1992

Sharon E. Rosenkoetter, Ph.D
Bridging Early Services Transition Project
Associated Colleges of Central Kansas
105 E. Kansas, McPherson, KS 67460
316 - 241-7754

Susan K. Wanska, Ph.D.
Department of Human Development and Family Studies
Justin Hall
Kansas State University
Manhattan, KS 66502
913 - 532-5510

* * *

What screening is -- and isn't

Definition: The application of quick and easily administered measures to a large population to differentiate those who may exhibit a target condition from those who are unlikely to evidence it

History

Originated in public health in the public schools during the 1920's to locate diseased children in order to limit the spread of infection

Was first applied to locate "physical and mental defects" in 1967 when P.L. 90-248 established EPSDT under Medicaid and began to require developmental screening

Rapidly expanded to Head Start, Home Start, the Parent-Child Center Program, and various disabilities services

Became institutionalized when P.L. 94-142 required all LEAs to identify, locate, and evaluate all children with disabilities residing in their jurisdictions -- soon came to be known as Child Find

Initially many screening programs used unvalidated measures; still occurs in many places this is extremely unfortunate, since screening is inherently a decision-making procedure, and important decisions should be made upon the basis of proven measures

Medical model continues to be very influential in screening, though other alternatives are available

Rationale for screening

1. To locate children with disabilities/developmental delays in order to provide them with intervention services

EC 301364

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Sharon E. Rosenkoetter
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

2. To minimize or eliminate existing problems and prevent development of secondary handicapping conditions
3. To provide parents with objective information about their children's developmental status
4. To support the family system of young children with potential disability through referral to appropriate resources
5. To provide a cohesive system for screening, evaluation, and intervention for eligible children

Requirements for effective screening measures

Reliability (consistency)

Test-retest reliability at .80 if used for individual decisions (van Doornick, 1974; Lehr, Ysseldyke, & Thurlow, 1987)

Small standard error of measurement on whole test and each subtest

Inter-examiner reliability of at least .80 (van Doornick, 1974; Lehr et al., 1987) -- depends upon clearly written manual and/or extensive training and monitoring of administrators

Validity (evidence that the procedure does what it claims to do)

Content relevance and coverage -- evidence that the task demands are related to the condition under discussion (e.g., being 3 years old, attending kindergarten) and that they survey the area

Criterion validity -- evidence that the procedure identifies (concurrent validity) or predicts (predictive validity); for group data, a correlation coefficient may be an adequate indicator, but for individual decision making (as in screening), a classification matrix is essential (See Attachment #1)

Construct validity -- evidence of theoretical basis that is claimed. While most developmental screening tests do not have an overall theoretical basis, many do claim that subsections evidence performance in different domains (and often these do not factor out into different skill areas as claimed)

Fairness (freedom from bias) -- evidence that the norm group for the measure is similar in language, culture, race, socio-economic status, and geography to the group which will receive the measure; lack of experience with testing or with school-like situations may also raise fairness issues in the screening of young children

An ethical, rather than quantitative, issue

"If the measure has more validity (concurrent or predictive) for one group rather than another, then it is biased" (Apling & Bryk, 1980)

Utility in the particular local situation

Procedures acceptable for local mores
Inexpensive, brief, simple to administer, easy to interpret
Acceptability for use with young children (see Attachment #2)
Individually administered
Attractive tasks in varied formats
Characteristics of young children which affect screening

Guidelines for best practice in screening (see Attachment #3)

Models for preschool screening
Mass screening or selective screening
Frequency of screening
Domains of coverage
Sources of information
Decision models for referral
Additive approach
Multiple cutoffs approach
Single stage or two stage
Variations on the models
Procedures and instruments
Tests (see Attachment #4)
Securing information from parents
Teacher input
Direct observation

Evaluation of screening procedures

Case study of an effective multi-faceted local preschool screening program (see Attachment #6)

What screening is -- and isn't -- revisited
Screening is **not** the basis for decisions about eligibility, placement, or category for special education service
Screening is **not** the basis for a decision to "wait a year" before starting kindergarten
Screening is **not** the basis for placement in a differentiated kindergarten class for "immature" children
Screening is **not** a procedure to determine instructional objectives or strengths/needs; a curriculum-referenced procedure is required for that purpose
Screening **is** the application of quick and easily administered measures to a large population to differentiate those who may exhibit a target condition from those who are unlikely to evidence it
Screening **is** an early step in a linked system to locate children with special needs and provide individually appropriate and age appropriate intervention

services to them and their families

Position statements (see list in packet & reprint in Attachment #5)

• • •

Attachment #1

Classificational Analysis To Determine a Procedure's Validity for Individual Decision-making

A classificational analysis shows the relationship between performance on the screening test and the criterion measure for each individual subject, as well as patterns of performance for all subjects. In order to determine a procedure's accuracy in identifying or predicting for individuals, it is necessary to do a classificational analysis in a 2 X 2 matrix, as demonstrated below:

		Performance on Criterion Measure (Concurrent or Subsequent) (e.g., eligibility for special education)	
		Positive	Negative
Performance on Screening Measure (e.g., DIAL-R)	Positive	A True Positive	B False Positive
	Negative	C False Negative	D True Negative

If A is very small and D is very large, as is true with the incidence of most disabilities, then a procedure might identify none of the children who actually have the disability and still have a high correlation coefficient. The literature contains numerous examples of procedures which identified fewer than 40% of the actual cases and still had high correlation coefficients (Meisels, 1989; Satz & Fletcher, 1979).

Prevalence rate, or base rate, for a condition is computed as follows:

$$\frac{A + C}{A + B + C + D} \times 100$$

Administrative decision-making about screening procedures requires knowledge about the pattern of misses. B and C above are both misses; B includes children

who erroneously screened positive for the condition, and C includes children who erroneously screened negative for the condition. Generally, in screening it is preferable for B to be higher than C, since only children in B will be followed up and evaluated for eligibility for early intervention services (Salvia & Ysseldyke, 1988). However, since all children in B must receive followup comprehensive evaluation, there may be financial or personnel reasons for reducing B; likewise, more families will become concerned about possible problems of their children if B is kept large. In any event, B and C are closely related to one another; if the screening procedure's cutoff is raised, B will increase while C decreases, but if the procedure's cutoff is lowered, C will increase while B decreases. Screening administrators must weigh the advantages and disadvantages of various error ratios; the decision as to cutoff point is a value-based one, depending upon local philosophy and resources (for an excellent example of local research on this issue, see Lindeman, Goodstein, Sachs, & Young, 1984). Standards for Educational and Psychological Testing (AERA, 1985) requires test publishers to justify their recommended cutoff score.

In order to determine the validity of a screening test for use in early identification, it is necessary to look at two types of formulas from the above matrix: horizontal and vertical.

Horizontal formulas look at "hit rates," correct identifications.

$\frac{A}{A + B} \times 100$ = % of "hits" from all children screened positive for the condition.

$\frac{D}{C + D} \times 100$ = % of "hits" from all children screened negative for the condition.

$\frac{A + D}{A + B + C + D}$ = % of "hits" from the total screening population. This is a measure of the overall percentage of accuracy in identification. Again, this frequently-cited "hit rate" does not reveal the pattern of misses or the relationship between correct negatives and correct positives. Subtracting the percentage of "hits" from 100% reveals the percentage of children that received false labels in order to locate those who were truly positive or truly negative.

Vertical formulas look at how successfully the screening procedure located the children who have (identification/diagnosis) or will have (prediction) the condition and those who do/will not.

$\frac{A}{A + C} \times 100$ = % of correct identifications of children who actually have the condition; this is called sensitivity, or, when reference tests are used to assign cases to cells, sometimes co-posivity.

$\frac{D}{B + D}$ X 100 = % of correct identifications of children who actually do not have the condition; this is called specificity, or, when reference tests are used to assign cases to cells, sometimes co-negativity.

It is essential to have the data from vertical analysis of the classification matrix to determine whether the screening procedure is accomplishing its purpose. Lichtenstein (1981) demonstrated that two frequently-used screening tests (i.e., the Denver Developmental Screening Test and Developmental Indicators for the Assessment of Learning) had the same correlation coefficients but disagreed totally in assigning children to high risk categories. Meisels (1989) suggested that both sensitivity and specificity, the factors computed in vertical analysis, should be at least 80% before a screening instrument is used to make individual decisions about children. Hobbs (1975) and McCall (1982) argued that screening procedures must reliably identify a significant number of children whose difficulties were not previously known; though no minimum percentages were given, this is another way of arguing for high sensitivity.

Summary. Criterion-relatedness is extremely important. No screening instrument should be evaluated for application to individuals without use of a classificational analysis.

* * *

Attachment #2

Characteristics of Young Children which Influence Screening Outcomes

Little or no experience with testing or school-like situations
Wariness of strangers
Separation anxieties at certain ages
Short attention span
Highly distractible
May focus on stimuli adults would not
Dramatic developmental changes within short periods of time
Tremendous variability in normal development within one child across various developmental domains, as well across children
Lack of motivation to please the examiner
Hour by hour variability in performance due to tiredness, hunger, etc.
May fear new places
Curious and eager to explore

Attachment #3

Guidelines for Best Practice in Preschool Screening

Screening and assessment should be considered an integral part of the total delivery of services -- not just a means of identification and measurement.

It is essential to use proven measures for screening, which is inherently a decision-making process: it is inappropriate to base potentially life-changing decisions about young children upon unproven measures.

Developmental screening should occur periodically: a child who does not exhibit a problem at one time may do so a year or two later due to changing environmental circumstances or to the difference in skills tested at various ages.

Multiple sources of information should be used to shape screening decisions.

Development should be viewed as a range of functioning, not as a fixed point.

Not all developmental areas are equally important indicators of current or future functioning.

Family information is critical to the screening process; it will help to determine whether or not to initiate more in-depth assessment. Sensitivity is essential in interviewing families.

Parents' right to full information and informed consent at all phases of the screening process must be honored.

According to the IDEA Amendments of 1991, the family has the right to direct the focus of the screening for children 0-2.

Confidentiality is essential. Parents, of course, have access to all screening results and reports for their child.

Processes, procedures, and instruments designed for screening should be used for that purpose and not for other purposes.

The setting and procedures for the screening should minimize fear, anxiety, and uproar.

Screening procedures must have demonstrated reliability, validity, and fairness for their intended use.

The more relevant and familiar the screening tasks and setting are to the child and to the family, the more likely it is that the results will be valid.

All aspects of the screening process must be culturally sensitive.

Examiners must be appropriately trained and monitored.

Developmental screening should be viewed as only one path to more in- depth assessment. Passing a screening test should not preclude further evaluation for intervention services if other risk factors suggest a need for it.

SUMMARY OF SELECTED SCREENING INSTRUMENTS

Battelle Developmental Inventory (BDI) Screening Test (1984)

Age Range: 6 months to 8 years

Description: Consists of 96 items from the Battelle Developmental Inventory; Items are included from the Personal-Social, Adaptive, Motor, Communication, and Cognitive domains.

Score types. Separate scores can be obtained for the total screening, personal-social, adaptive, cognition, receptive communication, expressive communication, total communication, gross motor, fine motor, and total motor components. Cutoff points are provided at 1.0, 1.28, 1.5, 1.65, and 2.33 standard deviations below the mean. Cutoff points are not available for birth to 0;5 age group.

Administration time. 10-30 minutes depending on age of the child, according to the manual; McLean et.al. (1987) reported an average time of 44.5 minutes.

Technical Qualities:

Standardization. The norms used to score the Screening Test are derived from the full scale Battelle Developmental Inventory; a separate standardization of the Screening Test has not been done.

Reliability. The authors report no reliability studies on the Screening Test in the manual. Reliability studies are reported for the BDI. McLean et.al. (1987) reported an interrater agreement of 95%.

Validity. Correlations between BDI Screening Test and full BDI are above .90 for all domains and the total score. Total BDI Screening Test scores were 90% sensitive and 80% specific to any handicapping condition. BDI age equivalents were found to be significantly lower than those derived from diagnostic measures, suggesting that states using percentage of delay scores should compute these only when children have already failed, based on cutoff scores, to prevent overidentification.

Out of 35 children with no prior indication of exceptionality, 22 were identified by the BDI Screening Test as needing further assessment. Of the 22 who received follow-up testing with the BDI, 19 were found to be performing above a cutoff of -2.0 SD below the mean and 3 were referred for more extensive assessments (McLean et.al., 1987).

Comments: The BDI Screening Test has several problems. Screening time may be longer than preferred. Materials must be assembled separately from the full BDI for ease of administration and materials are not readily portable. Instructions for scoring are not well presented. The BDI Screening Test has not been tested adequately and, therefore, should be used with caution (McLean et.al., 1987).

Denver Developmental Screening Test - II (DDST II) (1990)

Age Range: Birth - 6 years

Description: The DDST II consists of 125 items in 4 domains: Personal-social, Fine Motor-Adaptive, Language, and Gross Motor. Measurement includes parent report, observation, and direct elicitation.

Score types. Normal (no delays and a maximum of 1 caution), Abnormal (2 or more delays), Questionable (1 delay and/or 2 or cautions), Untestable (interpret depending on the number of refusals that would be delays or cautions if scored as failures).

Administration time. About 20 minutes

Technical Qualities:

Standardization. The DDST II was standardized in 1988 using a Colorado population. The sample, composed of 2096 children (1039 Denver, 1057 non-Denver), was recruited based on age, race, place of residence, and mother's education level.

Reliability. Only item reliability is reported in the manual. Interrater reliability for items ranges from 83% to 100% with a mean of 99.7%. Test-retest reliability for items ranges from 42.9% to 100% with a mean of 89%.

Validity. The manual reports no validity studies for the DDST II. The authors state, "The validity of the Denver II rests upon its standardization; i.e., it simply presents the age at which children in this sample are able to do a variety of tasks" (Frankenburg, et.al., 1990, p. 18).

Comments: The DDST II is brief and easy to administer. Because of the limited standardization sample and the lack of validity studies, the results of this test must be used with caution.

Developmental Indicators for the Assessment of Learning - Revised (DIAL-R) (1983, 1990)

Age Range: 2 - 6 years

Description: Station test with 3 domains: Motor (fine, gross); Concepts; and Language.

Score types. Item scores are converted to scale scores for age intervals of 3 months. Scale scores convert to cutoff scores for each of the 3 areas and for the total test.

Administration time. 20-30 minutes

Technical Qualities:

Standardization. National standardization with 2447 children (1358 white, 1089 nonwhite) completed in 1983; Stratification variables included age, sex, ethnicity, geographic region, and size of community. Supplemental data included language other than English spoken in the home and parents' level of education. Of the 2447 children in the standardization, the scores of 1861 children matched to 1980 census data ethnic breakdowns, were used to establish the cutoff scores in the 1983 manual. The norms in the 1990 manual are based on 1990 census projections and are derived by weighting all groups in the original standardization sample to match the 1990 census projection figures. The 1990 manual presents a renorming, not a standardization (Cohn, 1992).

Reliability. Correlations for test-retest reliability for total score range from .61 to .99. Interrater reliability for DIAL-R not found.

Validity. Most of the validity studies reported in the manual were conducted prior to the 1990 edition. When the standardization data were reanalyzed in 1990, test results for 123 children from the original standardization of the DIAL were compared to performance on the Stanford-Binet, using the 1990 DIAL-R norms. Using a cutoff of 1.5 SD for the DIAL-R total score, the authors found a sensitivity rate of 70.0% and a specificity rate of 90.3%. The overreferral rate was 8.9% and the underreferral rate was 2.4%. Using the 1.5 SD cutoff, the DIAL-R total score was not sensitive in identifying potentially advanced children (6.7%).

Comments: The team-based approach to this instrument makes it potentially expensive to use in terms of personnel and space required. The choice of different cutoffs for decision-making is useful for local areas. The 1990 edition, because of its renorming, is likely to identify more children as at risk and fewer children as potentially gifted than did the 1983 edition.

Developmental Profile II (DP II) (1972, 1986)

Age Range: Birth to 9 1/2 years

Description: Five domains: Physical (Motor), Self-Help, Social, Academic (Cognitive), and Communication; Can be administered through parent interview, combination interview and direct testing, or self-interview conducted by the teacher.

Score types. With the exception of the Academic Scale, raw scores can be converted only into age-equivalent scores. Neither percentile ranks nor standard scores are available. The Academic Scale score can also be converted into an IQ equivalency score.

Administration time. 20-40 minutes

Technical Qualities:

Standardization. Large (3008) normative sample but 91% of the sample came from Indiana and 9% came from Washington; 89% of the sample was from large cities.

Reliability. Reliability data are inadequate.

Validity. DP II correlates highly with several diagnostic measures of intelligence, achievement, language, and fine motor skills. One study reported that the DP II could be made more sensitive if children in the "borderline range" as well as in the "significant delay" range were referred; however, this does increase the number of overreferrals.

Comments: Some suggest the DP II be used only as a screening tool. It's considered technically inadequate given the many problems related to standardization, reliability, and validity.

Early Screening Inventory - Second Edition (ESI) (1988)

Age Range: 4 years - 5 years 11 months

Description: The ESI consists of 30 items related to 3 general areas of development: visual-motor/adaptive functioning, language and cognition, and gross motor/body awareness. Provides for parent input through a parent questionnaire. Other components include a physical examination and vision and hearing screening.

Score types. OK (higher than 1 SD below the mean), Rescreen (between 1 and 2 SDs below the mean), and Refer (lower than 2 SDs below the mean) based on the total ESI score by age.

Administration time. 15-20 minutes

Technical Qualities:

Standardization. Cutoff scores were established on the basis of a randomly selected sample of 465 public school children living in a middle class suburban community. All the children were white and spoke English. A large national standardization is in progress.

Reliability. Correlations for interobserver and test-retest reliability are reported to be .91 and .82-.91, respectively.

Validity. Using the McCarthy Scales of Children's Abilities as the criterion, the ESI has a sensitivity rate of 87% to 92% and a specificity rate of 90% to 95%. Using first grade school performance as a criterion, the ESI has a predictive sensitivity rate of 92% and a predictive specificity rate of 72%.

Comments: The ESI is an easily learned, brief screening test. Results from several reliability and validity studies show that the ESI predicts school performance with moderate to excellent accuracy through the second grade. The predictive sensitivity is

somewhat higher than its specificity. It overrefers more than it underrefers. Because the cutoff scores were based on a sample of middle class, Caucasian children, caution should be used in applying ESI cutoff scores to other populations. (Meisels, 1985)

Early Screening Profiles (ESP) (1990)

Age Range: 2 years - 6 years 11 months

Description: The ESP consists of 3 Profiles and 4 Surveys. The 3 Profiles are Cognitive/Language, Motor (Gross and Fine), and Self-Help/Social. The latter profile is completed by the child's parent or teacher. The 4 Surveys include Articulation, Behavior, Health History, and Home.

Score types. Two levels of scoring: Level I yields Screening Indexes based on the standard deviation or on percentiles. Cutoff points are determined by the user. Level II yields more in-depth information: standard scores, percentile ranks, and age equivalents.

Administration time. 15-30 minutes

Technical Qualities:

Standardization. National standardization, based on 1990 census estimates, was conducted in 1987-88; 1149 children (797 white, 197 black, 116 Hispanic, 39 were Native Americans, Asians, Alaskan Natives, and Pacific Islanders). Stratification variables were age, sex, ethnicity, geographic region, size of community, and parental education. Children with disabilities weren't systematically sampled but weren't excluded if they could respond adequately to items.

Reliability. Depending on the time interval between testings, test-retest reliability for the total score ranges from .73 to .83. No interrater reliability information was found for the total score.

Validity. The manual presents several predictive and concurrent validity studies using a number of criterion measures. Concurrent sensitivity rates for the total screening: Cognitive/Language + Motor + Self-Help/Social: Parent (53% to 92%); Cognitive/Language + Motor + Self-Help/Social: Teacher (61% to 78%). Concurrent specificity rates for the total screening: Cognitive/Language + Motor + Self-Help/Social: Parent (77% to 86%); Cognitive/Language + Motor + Self-Help/Social: Teacher (65% to 81%).

Comments: ESP was extensively researched prior to publication. It appears to lack sensitivity. ESP gathers data for multiple domains from multiple sources and settings. Parental input is recognized as an important factor in the screening process. Cutoff points can be determined locally. Easy to administer and attractive to children. Because each Profile is scored separately, users can opt to use the sections most relevant to their needs.

McCarthy Screening Test (MST) (1978)

Age Range: 4 - 6 1/2 years

Description: Consists of 6 subtests taken from the McCarthy Scales of Children's Abilities: Right-Left Orientation; Verbal Memory Part I; Draw-A-Design; Numerical Memory, Parts I and II; Conceptual Grouping; and Leg Coordination.

Score types. Subtest scores are used to identify a child as "at risk" or "not at risk," based on how many of the 6 subtests are failed at either the 10th, 20th, or 30th percentile rank relative to the MSCA standardization sample. Does not yield a total score.

Administration time. Approximately 20 minutes

Technical Qualities:

Standardization. No separate standardization for the MST; Normed using the standardization data from the MSCA.

Reliability. Reliability data reported by the author derived from evaluation of the MSCA; Others report a test-retest reliability of .80.

Validity. Using MSCA as a criterion, sensitivity is 67%-72% and specificity is 76%.

Comments: MST does not fare as well as another short form of the MSCA developed by Kaufman. Kaufman's short form has a group of 6 subtests which is child-oriented, has similar subtest proportions as the GCI that correlated well with the total MSCA GCI, is easy to administer and score, does not differentiate significantly by sex or race, and has sound psychometric qualities.

Miller Assessment for Preschoolers (MAP) (1988)

Age Range: 2 years 9 months - 5 years 8 months

Description: Consists of 27 "core" items assessing sensory and motor abilities and cognitive skills on both verbal and nonverbal tasks. Optional informal supplemental observation forms can be used to give evaluators a clinical framework for getting more in-depth information on qualitative aspects of the child's performance.

Score types. Total score yields a percentile score; Refer (0-5%), Watch (6-25%), Developmentally average (26-99%)

Administration time. Approximately 30 minutes to administer and score

Technical Qualities:

Standardization. National standardization in 1979-80; 1204 normal children (86% white, 12% black, 2% other); An additional 90 children with preacademic problems were included. Stratification variables were age, geographic region, race, sex,

size of residential community, and SES (based on parental education, professional category, and family income).

Reliability. Interrater reliability of .97 and higher on the total test and on 4 of the 5 indices. Test-retest reliability of .80 or higher for the total test and on 4 of the 5 indices.

Validity. With 90 children identified with preacademic problems, the MAP identified 84% of the children, 3;8 and older. The percentage of successful identification with younger children was lower.

Comments: The MAP includes many items reflecting neurological and sensory-motor maturity that have not traditionally been included in screening tests. Because all the test items are weighted equally and because of the number of items per index, the total score is more reflective of neuromotor than cognitive abilities. The MAP is probably the most useful for identifying preschool children with mild to moderate preacademic problems. The test is short, carefully developed, and well standardized.

Minnesota Child Development Inventory (MCDI) (1972)

Age Range: 6 months - 6 years

Description: Eight domains: general development, gross motor, fine motor, expressive language, comprehension-conceptual (receptive language), situation comprehension, self-help, and personal-social; Parent report format.

Score types. A single cutoff score, corresponding to a 30% delay, stratified by age and gender, indicates whether referral is needed; Scores falling between 20-30% are considered borderline. Development of local cutoff scores is suggested.

Administration time. Different sources suggest 20-30 minutes or 30-50 minutes

Technical Qualities:

Standardization. Standardized on 796 white children from middle to upper income suburban Minneapolis. Development of local norms is suggested.

Reliability. Internal consistency for General Development scale for different ages ranges from .87-.93.

Validity. Sensitivity ranges from .77-.88. Specificity ranges from .65-.77. MCDI General Development scale found to correlate well with cognitive scales of the Bayley, Cattell, Binet, and MSCA and to be a valid predictor of later McCarthy scores.

Comments: The MCDI is a 320-item Parent Report Inventory. It overrefers children not at risk. Some question its use as a screening inventory. The MCDI is not appropriate for infants and toddlers under 12-15 months of age. Revisions of the MCDI are planned. Companion instruments to the MCDI are the Infant Development Inventory (0-15 months) and the Preschool Development Inventory (3-6 years).

Attachment #5

Recommendations from U.S. Department of Education. (1991). Preparing Young Children for Success: Guideposts for Achieving Our First National Goal -- An America 2000 Education Strategy

Expand efforts to screen children, particularly the disadvantaged, early on to prevent developmental delays. Identification of children who may not yet be enrolled in school is a difficult task. However, schools, public health, and social service agencies should develop standardized approaches for finding and serving such children. In doing so, states and local agencies should build on information networks developed under the Individuals with Disabilities Education Act (IDEA) that bring together personnel from hospitals, Medicaid programs, state and local social service agencies, as well as pediatricians to identify infants and toddlers in need of special education services.

In conducting assessments, steps must be taken to ensure equitable and appropriate treatment:

--Assessment should be made by multidisciplinary teams that include educational and medical specialists as well as parents.

--To determine the need for such services as health care, nutrition, dentistry, counseling, and transportation, screening should cover the family, not just the individual child.

--Assessment should be ongoing so as to reevaluate the changing status of children diagnosed as having special needs.

--Determinations of educational and social need should not be based on social class, race, ethnicity, or gender.

Appropriate assessment: Use ongoing observations by teachers and parents to assess children's progress in cognitive, social, emotional, and physical domains. Systematic observations, portfolios of children's work, and reports from parents constitute appropriate assessment measures for young children. Training and time should be afforded teachers so that they can learn multiple assessment techniques and use the results to improve instruction.

Restrict use of standardized testing to identifying special needs and personalizing instruction. Given the rate at which young children grow, their inexperience with test taking, short attention spans, and the wide range of the early experiences, such test should not be used to determine their school entry or retention. Further,

the very traits we are most interested in measuring may bear only a moderate relationship to the skills that children will later develop.

School entry criteria: Base eligibility for school entry on chronological age; children should not be excluded from school on the basis of unfavorable cognitive, small motor, social, and/or emotional developmental assessments. We urge a halt to the practice of using standardized testing to screen out young children who are not academically "ready" for school entry or to place them in differentiated kindergarten programs. Schools have a responsibility to adapt their curricula to the capabilities of entering children, not to expect children to adapt to inappropriate school expectations and practices.

Attachment #6

A Longitudinal Study of Prekindergarten Transition Screening: Teacher Report and Traditional Testing (summary of article submitted for publication)

Sharon Rosenkoetter
Department of Special Education
Associated Colleges of Central Kansas
105 East Kansas Avenue
McPherson, Kansas 67460
316/241-7754

Lawrence Rosenkoetter
Department of Psychology
Bethany College
Lindsborg, Kansas 67456
913/227-3311

RATIONALE

Several professional organizations have recently issued position statements urging school administrators to use multiple measures in evaluating the progress of young children (National Association for the Education of Young Children, 1988; National Association of Early Childhood Specialists in State Departments of Education, 1987; National Association of Elementary School Principals, 1990). Such an approach frequently includes using early childhood teachers' and parents' appraisals of child development along with standardized tests to make educational decisions.

This project studied approximately 95% of the prekindergarten children without identified disabilities in one Midwestern county to determine whether preschool teacher appraisal, or multidisciplinary screening by a professional team, or both, best predicted learning problems two years later. The research was intended to evaluate empirically the policy recommendations of the national professional organizations as well as to assess the effectiveness of local practices in locating children eligible for special education services.

AIMS

1. To determine prior to children's kindergarten entry whether a preschool teacher checklist, an extensive multidisciplinary screening procedure, or a combination of the two, best predicts problematic school performance two years later, as measured by a teacher checklist and child placement information. This information was desired in order to locate children whose development might be enhanced by evaluation for special education services prior to kindergarten entry.

2. To analyze the role of child gender in the above predictions.

3. To learn which skill area deficits are most predictive of later school difficulties.

4. To demonstrate a procedure for evaluating local Child Find screening efforts.

METHODS, PROCEDURES, AND SUBJECTS

Data were collected on four different occasions, as indicated below:

#1 JANUARY

13 Preschool/Daycare Teachers

Complete 11 item checklist (see Attachment 1), 1-5 Likert scale

Subjects include 214 children (every one in a Midwest county who is enrolled at an early childhood center, eligible for kindergarten next fall, but not currently receiving special education services)

#2 FEBRUARY-MARCH

Prekindergarten Screening

1 1/2 hour, multidisciplinary screening--separate subtests for cognition, gross motor, fine motor, communication, and behavior, plus health and behavioral information from parents

Subjects include 288 children (182 from above, 106 new; none currently receiving special education services). This is approximately 95% of all the children in the county eligible to attend kindergarten next fall

#3 APRIL, TWO YEARS LATER

32 Present Teachers (kindergarten or first grade)

Complete same 11 item checklist as used by preschool teachers plus 3 additional questions about children's current placement and services (see Attachment 1)

Subjects include 258 children from one or both previous groups (149 from both groups 1 and 2; 24 from only group 1; 85 from only group 2) -- all that could be located. This totals 81% of the initial populations

#4 MAY

School Records

Is the child currently receiving special education services? If so, under what category?

258 children from group 3--all that can be located from the initial populations

RESULTS AND CONCLUSIONS

1. At the time of follow-up, 37 boys and 15 girls were judged by their teachers to be experiencing difficulty as measured by lack of readiness for next grade, consideration by the assessment team, placement in special services, recommendation for retention, and/or overall weak progress.

2. The 13 preschool teachers were hesitant to rate a child low; according to their responses, no teacher was "dissatisfied" (a rating of 1) with the progress of any child. Nevertheless, teacher ratings predicted 7% of the variance in problematic performance ratings two years later for all children -- 19% for girls and 2% for boys. Classification analyses appear in Figure 1. A lower cut-off on the teacher checklist resulted in more accurate predictions.

3. The school systems had a stated preference to "miss" children rather than to alarm too many parents by recommending an evaluation of their child for possible special education services. Nevertheless, the screening team predicted 13% of the variance for all children -- 19% for girls and 10% for boys. A classification analysis appears in Figure 2.

4. Far more powerful in predicting later school problems than either the teacher checklist alone or the multidisciplinary screening alone was a combination of the two. Merging the two sources of information resulted in a prediction of 20% of the variance for all children -- 28% for girls and 13% for boys. A classification analysis appears in Figure 3. The combined measurers were more accurate in predicting learning problems than either measure alone.

5. Both teacher appraisal and screening predicted markedly better for girls than for boys. Teachers (both preschool and primary grades) and screeners rated girls as a group higher on every single performance variable than they did boys. On 34 of 37 performance variables the standard deviation for girls was smaller than that for boys.

6. All preschool teacher ratings except self care were significantly correlated with teacher ratings of performance on the same variable two years later; correlations ranged from .27 to .46.

7. Screening subtests typically correlated only slightly with teacher ratings two years later, but the team recommendation for evaluation for special education correlated .37 with learning problems two years later.

8. Children presently in kindergarten, that is those who had waited out a year before starting school, scored similarly on both teacher checklists as their

first grade peers. Parent decision to "red shirt" these children apparently was based on factors which the teachers did not score on the checklist.

9. Older children were slightly more likely to have lower outcome scores (correlation .16, $p = < .05$).

10. Preschool teachers relied heavily on judgements about fine motor skills, independent work skills, academic skills, and direction following to evaluate a child's "readiness" for kindergarten. They were less concerned with communication skills, transitions, social development, classroom conduct, gross motor, or self care.

11. Kindergarten teachers relied on more variables in reaching their readiness conclusion: academic skills, direction following, independent work skills, social development, communication, and behavior during transitions between activities. They were less concerned with fine and gross motor, self care, and classroom conduct.

12. First grade teachers relied most heavily on academic skills, independent work skills, communication skills, direction following, fine motor, and transitions in determining their ratings. They were less concerned with classroom conduct, social development, gross motor skills, and self care skills.

13. In the screening, individual performance areas were less highly correlated with a team referral for further testing than was true for specified areas and "concern about readiness" on the teacher checklist; screeners relied most heavily on the following 6 subtests among 24 subtests in referring a child for evaluation for special education: verbal memory, auditory memory, motor skills, verbal concepts, vocabulary, and activity level. A score by a speech language pathologist suggesting difficulties in any of eight communication areas also resulted in an automatic referral.

14. Children who did not attend preschool/daycare were significantly more likely to experience learning problems two years later ($p < .05$), but this factor accounted for less than 2% of the variance in problematic outcomes.

15. Eleven of 13 preschool/daycare teachers stated that they appreciated the opportunity to offer their observations on children's present level of functioning and would like to do so in the future. Some would like to see their opinions shared with screeners in advance; others like having two independent judgements of child progress (which most often agree with one another) to inform parents about their child's development. Despite the fact that the purpose of this screening was to locate children with possible disabilities and not to track children into classes according to skill levels, two teachers did not wish to see children "labeled" prior to kindergarten entry. Thus they opposed any action by preschool or public school to assess children in advance of kindergarten, in spite of the fact that Child Find is

required by law.

16. Teachers reported that completing the checklist took from 2 to 30 minutes per child; some of this time was spent in discussing the child with colleagues before completing the form. Though no additional comments were encouraged, many preschool teachers (but no kindergarten or first grade teachers) wrote comments on the checklist. Most preschool teachers (but, again, no kindergarten or first grade teachers) said they shared the checklist and their responses to it with parents.

FIGURE 1

PRESCHOOL TEACHER EVALUATIONS COMPARED
WITH OUTCOMES TWO YEARS LATER

OUTCOME

SATISFACTION WITH CHILD PROGRESS	+	-	
+ (4,5)	122 (72%)	17 (10%)	
- (2,3)	17 (9%)	14 (8%)	136/169 correct 81% hit rate

OUTCOME

SATISFACTION WITH CHILD PROGRESS	+	-	
+ (3,4,5)	129 (76%)	20 (12%)	
- (2)	9 (5%)	11 (8%)	140/169 correct 83% hit rate

FIGURE 2

SCREENING RECOMMENDATIONS COMPARED WITH
OUTCOMES TWO YEARS LATER

OUTCOME

SCREENING TEAM EVALUATION OF CHILD FUNCTIONING	+	-
+	175 (75%)	33 (14%)
-	10 (4%)	17 (7%)

192/235 correct

81% hit rate

FIGURE 3

PRESCHOOL TEACHER EVALUATION PLUS SCREENING RECOMMENDATIONS COMPARED WITH OUTCOMES TWO YEARS LATER

OUTCOME

	+	−	
Both +	104 (70%)	10 (7%)	
At least 1− (screening or teacher rating of 2-3 = low)	17 (11%)	16 (11%)	122/149 correct 82% hit rate

PRESCHOOL TEACHER CHECKLIST PLUS SCREENING

OUTCOME

	+	−	
Both +	104 (75%)	12 (8%)	
At least 1− (screening or teacher rating of 2 = low)	12 (8%)	16 (11%)	125/149 correct 84% hit rate

References

Apling, R., & Bryk, A. (1980). Policy paper: The predictive validity of early childhood variables. Cambridge, MA: Huron Institute.

Bailey, D.B., & Wolery, M. (1989). Developmental screening: Rationale, methods, and application. Infants and Young Children, 4, 1-10.

Close Conoley, J., & Kramer, J.J. (1989). The tenth mental measurements yearbook. Lincoln, NE: University of Nebraska Press.

Cohn, M.E. (1992). Screening measures. In E. Vazque Nutall, I. Romero, & J. Kalesnik (Eds.), Assessing and screening preschoolers (pp 83-98). Boston, MA: Allyn and Bacon.

Frankenburg, W.K., Dodds, J., Archer, P., Bresnick, B., Maschka, P., Edelman, N., & Shapiro, H. (1990). Denver II Technical Manual. Denver: Denver Developmental Materials, Inc.

Glascoe, F.P. (1991). Developmental screening: Rationale, methods and application. Infants and Young Children, 4, 1-10.

Harrington, R.G. (1984). Preschool screening: The school psychologist's perspective. The School Psychology Review, 13, 363-374.

Harrison, P.L., Kaufman, A.S., Kaufman, N.L., Bruinicks, R.H., Rynders, J., Ilmers, S., Sparrow, S., Cicchetti, D., & McCloskey, G. (1990). Early screening profiles. Circle Pines, MN: American Guidance Service.

Keyser, D.J., & Sweetland, R.C. (1984, 1985). Test critiques: Vols I & II. Kansas City, MO: Test Corporation of America.

Hobbs, N. (1975). The futures of children: Categories, labels, and their consequences. San Francisco: Jossey-Bass.

Lehr, C.A., Ysseldyke, J.E., & Thurlow, M.L. (1987). Assessment practices in model early childhood special education programs. Psychology in the Schools, 24, 390-399.

Lichtenstein, R. (1981). Comparative validity of two preschool screening tests: Correlational and classificational approaches. Journal of Learning Disabilities, 14, 68-72.

Lichtenstein, R., & Ireton, H. (1984). Preschool screening: Identifying young children with developmental and educational problems. Orlando, FL: Grune & Stratton.

Lindeman, D.P., Goodstein, H.A., Sachs, A., & Young, CC. (1984). An evaluation of the Yellow Brick Road Test through a full prediction-performance comparison matrix. Journal of School Psychology, 22, 111-117.

Mardell-Czudnowski, C., & Goldenberg, D.S. (1990). Developmental indicators for the assessment of learning-revised. Circle Pines, MN: American Guidance Service.

McLean, M., McCormick, K., Baird, S., & Mayfield, P. (1987). Concurrent validity of the Battelle developmental inventory. Diagnostique, 13, 10-20.

May, D.C., & Welch, E.L. (1984a). Developmental placement: Does it prevent future learning problems? Journal of Learning Disabilities, 17, 338-341.

May, D.C., & Welch, E.L. (1984b). The effects of developmental placement and early retention on children's later scores on standardized tests. Psychology in the Schools, 21, 381- 385.

McCall, R.B. (1982). A hard look at stimulating and predicting development: The case of bonding and screening. Pediatrics in Review, 3, 205-212.

Meisels, S.J. (1992). Early intervention: A matter of context. Zero to Three, 12, 1-6.

Meisels, S.J. (1989). Can developmental screening tests identify children who are developmentally at-risk? Pediatrics.

Meisels, S.J. (1987). Uses and abuses of developmental screening and school readiness testing. Young Children, 42, 4-6 & 68-73.

Meisels, S.J. (1985). Developmental screening in early childhood: A guide. Washington, DC: National Association for the Education of Young Children.

Meisels, S.J.. & Provence, S. (1989). Screening and assessment: Guidelines for identifying young disabled and developmentally vulnerable children and their families. Washington, DC: National Center for Clinical Infant Programs.

Mitchell, J.V. (1985). The ninth mental measurments yearbook. Lincoln, NE: University of Nebraska.

Newborg, J., Stock, J.R., Wnek, L., Guidubaldi, J., & Svinicki, J. (1984). Battelle developmental inventory screening test. Allen, TX: DLM-Teaching Resource.

Salvia, J., & Ysseldyke, J.E. (1988). Assessment in special and remedial education. Boston: Houghton Mifflin.

Satz, P., & Fletcher, J.M. (1979). Early screening tests: Some uses and abuses. Journal of Learning Disabilities, 12, 56-60.

Shepard, L.A., & Smith, M.L. (1986). Synthesis of research on school readiness and kindergarten retention. Educational Leadership, 44, 78-88.

Thurlow, M.L., O'Sullivan, P.J., & Ysseldyke, J.E. (1986). Early screening for special education: How accurate? Eductional Leadership, 44, 93-95.

Tramontana, M.G., Hooper, S.R., & Selzer, S.C. (1988). Research on the preschool prediction of later academic achievement: A review. Developmental Review, 8, 89-146.

van Doornick, W.J. (1974). Guidelines for identifying handicapped children. In S.P. Hersh & S. Rojcewicz (Eds.), Health care screening and developmental assessment. Rockville, MD: National Institute of Mental Health.

Ysseldyke, J.E., Thurlow, M.L., & O'Sullivan, P.J. (1987). The impact to screening and referral practices in early childhood special education: Policy considerations and research directions. Journal of Special Education, 21, 85-96.

Ysseldyke, J.E., Thurlow, M.L., O'Sullivan, P., & Bursaw, R.A. (1986). Current screening and diagnostic practices in a state offering free preschool screening since 1977: Implications for the field. Journal of Psychoeducational Assessment, 4, 191-201.

Position Statements

American Academy of Pediatrics. (1981). Taskforce on P.L. 94-142. Washington, DC: Author.

American Educational Research Association, American Psychological Association, & the National Council on Measurement in Education. (1985). Standards for eductional and psychological testing. Washington, DC: American Psychological Association.

Boyer, E. (1991). Ready to learn: A mandate for the nation. Princeton, NJ: The Carnegie Foundation for the Advancement of Teaching.

Bredekamp, S. (Ed.) Developmentally appropriate practice in early childhood programs serving children from birth through age 8. Washington, DC: National Association for the Education of Young Children.

Meisels, S.J., & Provence, S. (1989). Screening and assessment: Guidelines for identifying young disabled and developmentally vulnerable children and their families. Washington, DC: National Center for Clincial Infant Programs (taskforce established by the National Early Childhood Technical Assistance System and the U.S. Department of Education).

National Association for the Education of Young Children & National Association of Early Childhood Specialists in State Departments of Education. (1991). Guidelines for appropriate curriculum content and assessment in programs serving children ages 3 through 8. Young Children, 46, 21-38.

National Association for the Education of Young Children. (1988). NAEYC position statement on standardized testing of young children 3 through 8 years of age. Young Children, 43(3), 42- 47.

National Association of Early Childhood Specialists in State Departments of Education. (1987). Unacceptable trends in kindergarten entry and placement: A position statement. Columbia, SC: NAECS/SDE.

National Association of Elementary School Principals. (1990). Early childhood education and the elementary school principal: Standards for quality programs for young children. Alexandria, VA: Author.

National Association of State Boards of Education. (1988). Right from the start: The report of the NASBE Taskforce on early childhood education. Alexandria, VA: Author.

National Association of State Boards of Education. (1991). Caring communities: Supporting young children and families. Alexandria, VA: Author.

National Commission on Children. (1991). Beyond rhetoric: A new American agenda for children and families. Washington, DC: Author.

U.S. Department of Education. (1991). Preparing young children for success: Guideposts for achieving our first national goal. Washington, DC: Author.

Child's Name _____ School _____

Present Grade (Spring 1989) _____ Teacher _____

Please quickly (without deliberation or discussion) circle your first reaction to these questions. Thank you for your assistance.

CHILD PERFORMANCE AREA	Severe Problem	Mild Problem	Not a Problem or a Strength	Strength	Definite Strength
1. Academic Skills (Comprehension; memory; visual and auditory discrimination; reasoning; concepts such as letters, numbers, colors, shapes)	1	2	3	4	5
2. Independent Work Skills (Working with minimal teacher assistance; attending to task; completing tasks neatly; correctly, and on time)	1	2	3	4	5
3. Following Directions (Follows instructions to group; complies with individual directions; responds to directions within a reasonable period of time)	1	2	3	4	5
4. Transitions (Stays engaged in transition activities, e.g., clean-up, locating and returning materials when finished, staying in line when moving to next area)	1	2	3	4	5
5. Communication (Communicates needs and preferences; understands language; speaks intelligibly with few errors; expresses ideas; has adequate vocabulary)	1	2	3	4	5
6. Social Development (Interacts with children and adults; shares; plays appropriately; expresses affection, emotions, and feelings appropriately)	1	2	3	4	5
7. Fine Motor/Perceptual Motor (Copying; cutting; drawing; writing; pasting)	1	2	3	4	5
8. Gross Motor (Walking; running; hopping; balancing)	1	2	3	4	5
9. Classroom Conduct (Behaves appropriately)	1	2	3	4	5
10. Self-care (Toileting; washing; grooming; dressing; eating)	1	2	3	4	5
11. Expected Readiness - Next Grade (based on current school performance	1	2	3	4	5

Not on Preschool Teacher Checklist

12. Has this student ever been referred to the Building team? ___ yes ___ no. If so, for what possible problems?

13. Has this student ever been recommended for retention in grade? ___ yes ___ no. If so, due to what deficiencies?

14. Has this student ever received any special education services? ___ yes ___ no. If so, under what category of exceptionality?

15. Overall, how satisfied are you with this child's progress and adjustment in your classroom?

Dissatisfied	Somewhat Dissatisfied	Neither	Somewhat Satisfied	Satisfied
1	2	3	4	5